The Patent Office

PCT/US 05/00612

REC'D 29 MAR 2005
WIPO PCT

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Signed AR Stewart

Dated 21 January 2005

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

03JAN04 E864009-7 D00056
P01/7700 0.00-0400334.9 CHEQUE

The Patent Office

Cardiff Road
Newport
South Wales
NP9 1RH

THE PATENT OFFICE
- 8 JAN 2004
LONDON

- 8 JAN 2004

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

1. Your reference — P32735GB

2. Patent application number *(The Patent Office will fill in this part)* — 0400334.9

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*
   M-Wave Limited
   6 The Lays
   Goose Street
   Beckington
   Somerset BA11 6RS

   Patents ADP number *(if you know it)* — 7841869004

   If the applicant is a corporate body, give the country/state of its incorporation — England

4. Title of the invention

   Data Compression Using Matching Pursuits

5. Name of your agent *(if you have one)*

   "Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*
   Kilburn & Strode
   20 Red Lion Street
   London
   WC1R 4PJ

   Patents ADP number *(if you know it)* — 125001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*
   a) any applicant named in part 3 is not an inventor, or
   b) there is an inventor who is not named as an applicant, or
   c) any named applicant is a corporate body.
   See note (d))

   YES

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	23
Claim(s)	10
Abstract	1
Drawing(s)	6 + 6 SW

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	1
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Signature Kilburn & Strode

Date 08/01/2004

12. Name and daytime telephone number of person to contact in the United Kingdom

Michael N Maggs
Tel: 020 7539 4200

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

Data Compression Using Matching Pursuits

The present invention relates to data compression techniques, and in particular to techniques which use the matching pursuits algorithm. The invention is particularly although not exclusively applicable to the field of video and still image compression.

The transform known as matching pursuits was introduced by Mallat & Zang, in their paper "Matching Pursuits with Time-Frequency Dictionaries", IEEE *Transactions on Signal Processing, volume 41, 12 December 1993*. A significant amount of research has been carried out subsequent to publication of that paper with a view to applying the matching pursuits algorithm to the compression of video images and audio data, as exemplified by Neff & Zakhor, *"Very low bit rate video coding based on Matching Pursuits", IEEE Transactions on Circuits and Systems for Video Technology, volume 7, number 5, October 1997, pages 158-171*; see also their US patent US-A-5699121. While the transform has proved extremely effective, its practical application has been limited, primarily because it requires a large amount of computation and is therefore relatively slow. Conventional thinking is that the transform is unlikely to be useable in practical real-time video coding systems for some years to come, until the transform has been sufficiently optimised or hardware speeds sufficiently increased. One approach however, that has suggested using matching pursuits in the content of the encoding of speech is that described by Rezaiifar and Jafarkhani in their paper "*Wavelet Based Speech Coding Using Orthogonal Matching Pursuit*", *Proc. 29th Int. Conf. on Information Systems (CISS – 95), pp88-92, Mar 1995.*

The invention also dispels controversial thinking matching pursuit must always be computationally intensive.

According to one aspect of the invention there is provided a method of data compression comprising applying a transform to multi-dimensional data to generate a multi-dimensional transform data set, and coding the transform data set by applying one or more one-dimensional matching pursuits algorithms.

Preferably, a plurality of one-dimensional matching pursuits algorithms are used, each in a different scanning direction through the data. The scan directions may (but need not) be orthogonal. There may be a single one-dimensional matching pursuits algorithm per dimension of the data, or there may be fewer: in other words we may use one or more matching pursuits algorithms up to the number of dimensions of the data.

According to another aspect of the invention there is provided a method of data compression comprising:

(a) applying a transform to multi-dimensional data to generate a multi-dimensional transform data set;

(b) convolving the transform data set with each of a plurality of first one-dimensional basis functions to generate a corresponding plurality of convolved data sets;

(c) determining a location in a first direction across all the convolved data sets, and a first basis function, representative of a greatest magnitude;

(d) convolving the transform data at the said location with each of a plurality of second one-dimensional basis functions;

(e) determining a second basis function representative of a greatest magnitude;

(f) representing part of the transform data surrounding the said location with an atom derived from the first and second basis functions corresponding to the greatest determined magnitudes;

(g) subtracting the atom from the transform data set to create a new data set;

(h) repeatedly updating the convolved data sets by convolving any changed part of the transform data set with each of the plurality of first one-dimensional basis functions, and then re-applying steps (c) to (f); and

(i) outputting as quantized transform data coded versions of the atoms derived at step (f).

According to another aspect there is provided a method of data compression comprising:

(a) applying a transform to multi-dimensional data to generate a multi-dimensional transform data set;

(b) convolving the transform data set with each of a plurality of first one-dimensional basis functions to generate a corresponding plurality of convolved data sets;

(c) determining a first location in a first direction across all the convolved data sets, and a first basis function representative of a greatest magnitude; and representing part of the transform data surrounding the first location with a first atom derived from the first function corresponding to the greatest determined magnitude;

(d) subtracting the first atom from the transform data set to create a new data set;

(e) convolving the new data set with each of a plurality of second one-dimensional basis functions;

(f) determining a second location in a second direction across all the convolved data sets, and a second basis function representative of a greatest magnitude; and representing part of the new data set surrounding a second location with a second atom derived from the second function corresponding to the greatest determined magnitude;

(g) subtracting the second atom from the new data set to create a further new data set;

(h) repeating step (b) with the further new data set, and then re-applying steps (c) to (f); and

(i) outputting as quantized transform data coded versions of the atoms derived at steps (c) and (f).

According to another aspect there is provided a coder for data compression comprising means for applying a transform to time-varying data to generate a multi-dimensional transform data set, and a coder for coding the transform data set by applying a plurality of one-dimensional matching pursuits algorithms, one for each dimension.

According to another aspect there is provided a coder for data compression comprising:

(a) means for applying a transform to multi-dimensional data to generate a multi-dimensional transform data set;

(b) means for convolving the transform data set with each of a plurality of first one-dimensional basis functions to generate a corresponding plurality of convolved data sets;

(c) means for determining a location in a first direction across all the convolved data sets, and a first basis function representative of a greatest magnitude;

(d) means for convolving the transform data at the said location with each of a plurality of second one-dimensional basis functions;

(e) means for determining a second basis function representative of a greatest magnitude;

(f) means for representing part of the transform data surrounding the said location with an atom derived from the first and second basis functions corresponding to the greatest determined magnitudes;

(g) means for subtracting the atom from the transform data set to create a new data set;

(h) means for repeatedly updating the convolved data sets by convolving any changed part of the transform data set with each of the plurality of first one-dimensional basis functions; and

(i) means for outputting as quantized transform data coded versions of the derived atoms.

According to another aspect there is provided a coder for data compression comprising:

(a) means for applying a transform to multi-dimensional data to generate a multi-dimensional transform data set;

(b) means for convolving the transform data set with each of a plurality of first one-dimensional basis functions to generate a corresponding plurality of convolved data sets;

(c) means for determining a first location in a first direction across all the convolved data sets, and a first basis function representative of a greatest magnitude; and representing part of the transform data surrounding the first location with a first atom derived from the first function corresponding to the greatest determined magnitude;

(d) means for subtracting the first atom from the transform data set to create a new data set;

(e) means for convolving the new data set with each of a plurality of second one-dimensional basis functions;

(f) means for determining a second location in a second direction across all the convolved data sets, and a second basis function representative of a greatest magnitude; and representing part of the new data set surrounding a second location with a second atom derived from the second function corresponding to the greatest determined magnitude;

(g) means for subtracting the second atom from the new data set to create a further new data set;

(h) means for repeating step (b) with the further new data set, and then re-applying steps (c) to (f); and

(i) means for outputting as quantized transform data coded versions of the atoms derived at steps (c) and (f).

The invention further extends to a codec including a coder as previously described. It further extends to a computer program for carrying out the method described, and to a machine-readable data carrier which carries such a computer program.

In the preferred method, the transform consists of or includes a decorrelating transform and/or a frequency based transform.

In applying the matching pursuits algorithm, the mechanism for convolving the transform data set with each of the plurality of the bases is not critical. Typically, this may be achieved by calculating the inner product of each of the bases with every possible position (data point) in the transform data set. However, less accurate methods of locating the position may also be used. Likewise, the position where the inner product is greatest may be determined in any convenient way, for example by searching. Preferably, a small portion of the data around the relevant point is then represented by the basis function at that position multiplied by a coefficient which has the same sign as the selected inner product, and the square root of its magnitude.

The position having the greatest magnitude may be determined by taking the absolute magnitude (that is, relative to to zero). Alternatively, the position of greatest magnitude may be determined after the application of a function across the data which may represent a sensory or psychophysical model such as a psychoacoustic or psychovisual model representative of the perceptual importance of the data. The function map may, but need not, define threshold values which are subtracted from the data before the position of greatest magnitude is determined. Alternatively, the function map may be used as a multiplier to the data, or combined with it in some other way.

The method of the present invention may be used to compress both two-dimensional data (for example still images), as well as three-dimensional data (for example moving images with some compression in the time dimension). When three-dimensional data is to be compressed, a two-dimensional transform may be used, followed by three one-dimensional matching pursuits algorithms.

In one embodiment of the invention, the whole or part of the transform data set (for example a sub-band of the data set) may be scanned in the direction in which the data is most correlated, and one-dimensional matching pursuits may be applied to the data so scanned.

The invention may be put into practice in a number of ways and several specific embodiments will now be described, by way of example, with reference to the accompanying drawings in which:

Figure 1 illustrates a first embodiment of the present invention in which independent one-dimensional atoms are used;
Figure 2 illustrates a second embodiment in which two one-dimensional atoms are used to generate a two-dimensional atom;
Figure 3 illustrates a generic method of video coding;
Figure 4 shows in schematic form a video coder according to an embodiment of the present invention;
Figure 5 shows a decoder corresponding to the encoder of figure 2; and
Figure 6 illustrates an application of matching pursuits to the two-dimensional output of a wavelet transform.

Before describing the specific embodiments in detail, it may be worth summarising the operation of the matching pursuits transform. Specifically, we will summarise the way in which a 2D transform may be used to compress a 2D block of data, such as a still image.

Matching pursuits in the 2D case uses a library of 2D basis functions, typically normalized Gabor functions although other functions are equally possible. To encode the image, the transform forms the inner product of all the bases with every possible data point. This is equivalent to convolving the data with each and every basis function. Locally, wherever the basis function resembles the data, peaks will occur within the inner product. The results are then searched for the basis function that gives the inner product of largest magnitude: we can then represent a small portion of the data by the basis function at that position, multiplied by a coefficient which has the same sign as the selected inner product and the square root of its magnitude.

This gives what is known as an "atom". The code for the atom is the amplitude and the position within the data set (e.g. image), along with the number of the corresponding basis function.

The atom just found is then subtracted from the image, giving a modified image which represents the so-far uncoded portion of the data. This process is then iteratively repeated, to find additional atoms. At each iteration, a search is carried out for the basis function that gives the inner product of largest magnitude. Of course, it is necessary to update the convolution only where subtraction of the basis function has changed them.

The atoms found at each iteration are simply summed to create the encoded version of the data. When sufficient atoms have been found to represent the original image at some desired level of fidelity, the resultant list of atoms constitutes the compressed code. This code can be arranged and entropy coded, if required, to reduce its size.

The one-dimensional matching pursuits transform is similar, except of course that the code book consists of 1D rather than 2D functions. 1D matching pursuits has been applied as a transform to raw audio data with promising results, although, as mentioned above, the fact that the transform is computationally intensive has until now severely limited its usability in practical real-time applications.

The operation of the invention according to a first embodiment is shown schematically in figure 1. Here, some sort of transform (preferably a frequency transform) has been applied to a multidimensional data set (not shown) to create a multidimensional transform data set 10. For the sake of illustration only, the data set 10 is a two-dimensional data set having axes x,y.

To apply matching pursuits to this data set, the data is raster-scanned in the x direction, and the location 12 is determined at which one finds an atom of greatest magnitude. Position, amplitude and code book entry are recorded. The atom just found is then subtracted – typically after quantization of the amplitude - from the transform data set 10 to create a modified data set 10a. Next, the data set 10a is raster scanned in the y direction and the process repeated to find the best y atom at the further location 14 (which may not be the same as location 12).

To encode a two-dimensional image, this process is simply repeated, with the scans being taken alternately in the x and in the y directions. While the scans are preferably orthogonal, it is not essential that they be horizontal and vertical, and in some applications it may be preferable to use alternating raster scans which go diagonally.

This embodiment may easily be extended to higher dimensional data sets. For example, a three-dimensional data set may be encoded with alternating x,y,z scans, with an appropriate one-dimensional atom being selected at each of the scans. When the data set is representative of a video stream in x,y,t, the same approach may be used with the t axis being treated in exactly the same way as the z axis above. In other words, the t axis may be treated as if it were an independent spatial axis. Time-varying three-dimensional data in x,y,z,t may also be treated in a similar way, with repeated scans being made in x,y,z,t, x,y,z,t and so on. As with the two-dimensional case, while it is preferred that the scans are made in mutually orthogonal directions, it is not essential for the axes to be those previously mentioned. In some embodiments, raster scanning across diagonal planes may be preferred.

The code book used for each one-dimensional scan may be unique, or alternatively the same code book may be used for scans in more than one direction. It may be desirable for a first code book to be used for scans in the spatial dimensions, and for a second code book to be used for scans in the time dimension.

It is not essential for each raster scanned to be taken over the entirety of the data set to the encoded. Where desirable, the data set may be partitioned before scanning is undertaken, with the scans being carried out on each partition separately. The

partition may be the spatial partition, a temporal partition, a frequency partition, or any other type of partition that may be convenient according to the type of data being processed and the particular transform that is in use. 1D scanning may be done in different directions in each region.

An alternative and rather more sophisticated approach is shown in figure 2. Here, scanning is first carried out in the x direction 20 and the location 22 of the best-fitting x-atom 22 is determined as before. Next, an orthogonal scan 24 in the y direction is carried out, not across the whole data set but instead just locally in the region of the atom 22. At this location, the best y-atom is then selected. The x-atom and the y-atom together define a single (separable) two-dimensional transform. The amplitude is quantized, and the atom is subtracted from the data set to create a modified data set on which the procedure may be repeated. Because the second scan is taken in a direction orthogonal to the first, it will be understood that the location of the x-atom 22 becomes split up in the y-axis output stream as shown in the lower part of the figure at 22a, 22b, 22c. The "repair" of the data set following the subtraction therefore needs to be carried out at multiple places within the y-axis stream.

The same approach may be used in more than two dimensions. So, where the original data set is representative of a three-dimensional model in x,y,z, three separate one-dimensional scans may be used to generate a single 3D atom. Likewise, for encoding a video stream in x,y,t, the t-axis may simply be treated as a third spatial dimension to create a single 3D atom in x,y,t. Where a sequence of video images is to be encoded, a suitable transform is first applied, as discussed above, to create a three-dimensional transform data set (in space and time). Then, on the three-dimensional data set we carry out three matching pursuits separable

searches, each in a different orthogonal direction. So, first, we may for example carry out a raster scan through the data in the x direction, using a first one-dimensional code book, and look for the best matched location for each atom within that code book. Once the best x atom and location have been found, we then use matching pursuits again, but this time looking in the y direction only and using a separate y-code book. It is not necessary to scan the entire data set again, in the y direction, as we have found that in practice restricting the y search to a small area at or very near the previously-identified best location can still provide good results while substantially reducing computational overhead.

Once the best y atom has been located, the process is repeated, this time using a one-dimensional code book based upon the time dimension. As before, the t search may be restricted to an area at or close to the best locations previously found in the x and/or y searches. It has been found in practice that the best results can be obtained by using separate matching pursuits code books for the x, y and t searches. However, where appropriate, a common code book may be used either just for the x and y directions, or for all three directions.

Once all three one-dimensional atoms have been identified, an entire three-dimensional block around the preferred location can then be reconstructed. We recompute (using the inner product) to reconstruct the adjusted data (including the amplitude, since the "amplitude" of each individual direction is not sufficient, in itself, to calculate the actual model amplitude).

The modelled data is then quantized and subtracted from the original data. A note is made of the three one-dimensional atoms for future reference, and the entire process is then repeated on the reduced data set. This continues until the data set

has been reduced sufficiently, accordingly to the requirements of the particular application.

The same approach may be applicable to encode time-varying three-dimensional data, by means of single 4D atoms constructed from separable x,y,z and t atoms.

While the above sets out a variety of possible options, the most preferred implementation is as follows: after the multidimensional transform we simply scan the data in any desired 1D readout order. We then code the 1D scan with 1D Matching Pursuits.

This may be repeated by re-scanning in some other readout order (which may but need not be orthogonal to the first). Thus, we typically use one or more matching pursuits algorithms to a maximum of one per dimension of the data.

Where a wavelet transform has been used, the x code book may include atoms which define frequency, phase, attenuation, amplitude constant size; the y code book may define slew, attenuation, size and amplitude (frequency need not to be considered since that has already been decided); and the t code book may define time-slew, attenuation and size.

A convenient function that may be used to define the individual one dimensional atoms is the Gabor function. In the x direction, that is given by *f(x),* where:

$$F(x) = A\cos(\omega x + \Phi)\, e^{-\lambda t}$$

Similar Gabor functions G(y) and H(t) may be used for the y and t directions. It will be understood, of course, that the amplitude, phase shift and attenuation constants may be different in each of the three directions.

The three matches are not necessarily carried out in the order described above, and it may sometimes be sufficient or convenient to match in some other order, for example t, x, y, or x, t, y.

In the preferred embodiment, each 1D matching pursuit algorithm is implemented by calculating the inner product of each of the available bases with every considered position (data point) in the transform data set. The position where the inner product as the greatest absolute magnitude is then determined in any convenient way, for example by searching. As mentioned above, the second and third 1D matches do not require any searching in space at all, although in some circumstances it may be convenient to carry out a small 1D (or 2 or 3D) search in the vicinity of the 'best' location found by the previous matches Alternatively, instead of looking for the position having the greatest absolute magnitude (that is, relative to zero) some other measure of magnitude may be used instead. Where a model such as a psychoacoustic or a psychovisual model is in use, the position of greatest magnitude may be chosen with reference to that underlying model. One way of doing that is to apply the model as a weighting over the entire set of data before determining the position of greatest magnitude; alternatively, another approach is to view the model as defining thresholds, in which case the thresholds may be subtracted from the inner products following which the difference of greatest magnitude is sought. It will be understood, of course, that other approaches are possible. In general, a psychoacoustic, psychovisual or other model may be applied as a function map across the transform data set, with the position of

greatest magnitude being located across the transformed data as modified by the mapped function.

In any of the embodiments described above, instead of alternating between scanning directions an approach we call "agile scanning" may be used instead. We start by scanning separately in each possible direction, and determining the best possible atom in each of those directions. The amplitude (magnitude) of each of those atoms of is stored. Next, we repeatedly scan the channel (direction) of the highest magnitude until that channel generates an atom having a magnitude which is *less than* that of the stored magnitude of one of the other channels. Then, we switch to the channel which currently contains the atom of greatest magnitude and repeatedly scan that in the same way. We switch again when that channel no longer generates an atom of highest magnitude.

If the most recently scanned channel generates an atom of an identical magnitude to one which has already been found in another channel, we prefer that the channels should be switched. Alternatively, however, it would equally be possible never to switch channels in such a situation.

This approach is particularly efficient, since it allows the encoder to concentrate on obtaining "quick gains" in one channel, and automatically to switch to another channel as soon as it becomes optimal to do so. Since the rules used are causal, the state of the encoder can continually be tracked by corresponding rules set up within the decoder without the need to transfer any status bits.

The preferred embodiments, described above, provide for the first time the promise of a full 3D matching pursuits video coder that does not require the use of motion

vectors as in conventional motion compensated video coding. The need for motion vectors is effectively eliminated by the temporal aspect of the atoms. This has the effect of making fully-scalable video coding a real possibility for the first time.

Of course, it is by no means excluded that motion compensation may still be used when desired. In such an embodiment, applicable to both time-varying 2D and 3D data, once one or more spatial atoms have been determined, those atoms may then be copied into (or moved along) the time dimension by some prediction mechanism such as for example the use of motion vectors. This approach will now be discussed in more detail, with a view to illustrating how embodiments of the present invention may be incorporated within a motion-compensated codec.

To set the scene for these specific embodiments, we will next describe, briefly, some standard motion-compensated video compression techniques.

Video compression is divided into two basic categories: motion-compensated and non motion-compensated. When individual frames are compressed without reference to any other frames, the compression is described as "intra-coded". One of the advantages of intra-coded video is that there is no restriction on the editing which can be carried out on the image sequence. As a result, most digital video in the broadcasting industry is stored in this way at source. The intra-coding approach can be used in association with any of a large number of still image compression techniques such as, for example, the industry standard JPEG compression scheme. This approach is taken by the moving JPEG standard for video compression: JPEG compression is used for each of the individual frames, with each of the frames being handled independently and without reference to any other frame.

Video sequences are not, however, typically composed of a collection of entirely unrelated images, and greater compression can normally be obtained by taking account of the temporal redundancy in the video sequence. This involves a process known as inter-coded compression. With this approach, individual images in the output sequence may be defined with reference to changes that have occurred between that image and a previous image. Since the compressed data stream (sent across the video channel for reconstruction by the decoder) typically represents information taken from several frames at once, editing on the compressed data stream is not normally carried out because the quality is severely compromised.

Inter-coded compression is one of the compression techniques that is incorporated into the MPEG video compression standard.

A typical inter-coded compression scheme is shown schematically in Figure 3. In that Figure, the upper row 0 represents the original digitised video frames that are to be compressed, the second row C represents the compressed images, and the bottom row R the residuals.

In the scheme shown, selected original frames S are treated as still images, and are compressed by any convenient method to produce intra-frames 1. These frames are then used as reference frames to create predicted frames P. The contents of these frames are projected from one or more I frames - either forwards or backwards in the sequence. This is normally achieved by the use of motion vectors, associated with moving blocks within the image. Alternatively, the movement of specific physical objects within the image may be determined and predicted. Finally, the C sequence is completed by generating interpolated frames

B between the P and I frames. The original video sequence can then be approximated by the sequential frames of the C sequence, namely the I, B and P frames. In practice, further corrections normally have to be made if the end result is to appear reasonable. These further corrections are achieved by determining a residual frame R corresponding, in each case, to the difference between the original frame and the corresponding compressed frame. Residual frames may, but need not, be calculated for the intra frames. Accordingly, the residual frames marked X may sometimes be omitted.

In a practical embodiment, an encoder calculates the I frames from the these original frames labelled S in the diagram, and, from that, calculates the motion parameters (vectors) that are needed to define the P frames. The data stream transmitted from the encoder to the decoder thus includes the encoded I frames and the appropriate motion vectors enabling the decoder to construct the P frames. Information on the B frames is not sent, since those can be reconstructed by the decoder alone purely on the basis of the information within the I and P frames. In order to improve the final result, the data stream also includes the residual images, sent on a frame by frame basis. Since the residual image represents the difference between the original image and the compressed image, the encoder needs to have access to the sequence of compressed images. That is achieved by incorporating an additional decoder within the encoder.

The final data stream, as sent, therefore includes the full I frames, the motion vectors for the P frames and all of the residual frames possibly excluding those that are labelled X in Figure 1. Each residual image is typically compressed before transmission.

Numerous transforms, including matching pursuits, are known in the art for compressing the original S frames to produce the Intra frames. It has also been suggested, in the Neff and Zachor paper mentioned previously, that matching pursuits may be used to encode the residual images.

In contrast, in the preferred embodiment, the raw images are transformed by means of any standard transform, and the output of the transform is then *quantized* using the matching pursuits algorithm. The same applies to any residual images: instead of applying matching pursuits as a transform to the residual image, the residual image is instead first transformed using a standard transform, and the output of that transform is then *quantized* using matching pursuits. In both cases, the initial transform which operates on the data itself may for example be an FFT, a wavelet transform, a DCT or a lapped orthogonal transform. Other transforms could also be used.

Where motion vectors are to be used, the methods discussed above may be incorporated within a motion-compensated hardware or software encoder, as shown in Figure 4, although as previously mentioned motion vector compensation is not necessarily required at all in the present invention.

As shown in figure 4, frame by frame input is applied at an input 302, with the intra-frame data being passed to an intra-frame transform 304 and then to a matching pursuits coder or atom finder 303. The atom the amplitude is then quantized at 305. The inter-frame data is passed to a motion estimator 306 which provides a parametised motion description on line 308, this then being passed to a

motion compensator 310. The motion compensator outputs a predicted frame along a line 312 which is subtracted from the input frame to provide a residual frame 314 which is passed to a residual transform 316. The transform output is applied to a matching pursuits coder 309 and then to a quantizer 307 which outputs quantized codes to the output stream.

The motion description on line 308 is also passed to a motion description coder 320, which codes the description and outputs motion data on a line 322.

The output stream thus consists of coded intra-frame data, residual data and motion data.

The output stream is fed back to a reference decoder 324 which itself feeds back a reference frame (intra or inter) along lines 326, 328 respectively to the motion compensator and the motion estimator. In that way, the motion compensator and the motion estimator are always aware of exactly what has just been sent in the output stream. The reference decoder 324 may itself be a full decoder, for example as illustrated in Figure 5.

Generally, the motion vectors may be derived by comparing a successor frame with the decompressed previous frame; in the alternative, the original previous frame could be used. In either case, the residual frames are calculated as the difference between the predicted frame and the original successor frame. In a variation (not shown) of the embodiment, the frames being compared might be pre- to improve the motion vectors.

The output stream travels across a communications network and, at the other end, is decoded by a decoder which is shown schematically in Figure 5. The intra-information in the data stream is supplied to an intra-frame decoder 410, which provides decoded intra-frame information on a line 412. The inter-information is supplied to a bus 414. From that bus, the residual data is transmitted along a line 416 to a residual decoder 418. Simultaneously, the motion data is supplied along a line 420 to a motion compensator 422. The outputs from the residual decoder and the motion compensator are added together to provide a decoded inter-frame on line 423.

Reference frame information is fed back along a line 424 to the motion compensator, so that the motion compensator always has current details of both the output from and the input to the decoder.

It will be understood of course that the invention is not restricted to use with the type of motion-compensated video coder as shown in Figure 4: it may be used in any type of video coder, where the output from the main transform needs to be quantized.

This approach may be used not only for video compression, as previously described, but also for still image compression.

A further embodiment, in which the raw input data is for example representative of a still image, is shown in figure 6. Here, the input data/image 40 is first transformed/compressed in some way (e.g. by means of a wavelet transform 41), to produce a transformed image 42. That image is then quantized by means of a matching pursuits coder and quantizer 43 to produce the final coded output 44.

The wavelet transform 41 could be replaced with any other convenient compression transform such as an FFT a DCT, or a Lapped Orthogonal transform.

In the example shown in Figure 6, the image 40 undergoes a wavelet transform which splits the image up into several spatially-filtered sections or sub-bands 45, 46, 47, 48. Sections 46 and 47 have been highpass filtered in one direction and low pass filtered in another, which means that those two sub-bands are better decorrelated in one direction than in the other. It will be understood, of course, that a horizontal transform could be followed by a vertical, or vice versa. After raster scanning those sub-bands as indicated by the reference numerals 400, 401, a one-dimensional matching pursuits quantization is then used. A fairly small matching pursuits code book may be used, for each direction, since the finding of structure within the image at different scales has already been automatically carried out by the wavelet transform: that no longer needs to be carried out by the matching pursuits algorithm.

The use of 1D matching pursuits algorithms to quantize the output of a 2D transform is applicable not only to wavelet transforms but to any other 2D transforms which decorrelate better in one direction than in another (at least over part of the area of the output). Generally, the output of the transform may automatically divide the data up into a number of different partitions, and the partitions may then be individually scanned, each in a separate preferred direction.

It is envisaged that the matching pursuits algorithm may be applied to the output of any multi-dimensional decorrelating transform, preferably a frequency transform.

Claims:

1. A method of data compression comprising applying a transform to multi-dimensional data to generate a multi-dimensional transform data set, and coding the transform data set by applying a one or more one-dimensional matching pursuits algorithms.

2. A method of data compression comprising:

   (a) applying a transform to multi-dimensional data to generate a multi-dimensional transform data set;

   (b) convolving the transform data set with each of a plurality of first one-dimensional basis functions to generate a corresponding plurality of convolved data sets;

   (c) determining a location in a first direction across all the convolved data sets, and a first basis function, representative of a greatest magnitude;

   (d) representing part of the transform data surrounding the said location with an atom derived from the first and second basis functions corresponding to the greatest determined magnitudes;

   (e) subtracting the atom from the transform data set to create a new data set;

   (f) repeatedly updating the convolved data sets by convolving any changed part of the transform data set with each of the plurality of first one-dimensional basis functions, and then re-applying steps (c) and (d); and

(g) outputting as transform data coded versions of the atoms derived at step (d).

3. A method of data compression as claimed in claim 2 in which the coded version of each atom includes magnitude, position in transform data set and number of basis function.

4. A method of data compression as claimed in any one of the preceding claims in which the data to be compressed represents video image data.

5. A method of data compression as claimed in any one of claims 1 to 3 in which the data to be compressed represents a still image.

6. A method of data compression as claimed in any one of claims 1 to 3 in which the data to be compressed comprises residual images within a motion compensated video coder.

7. A method of data compression as claimed in any one of claims 1 to 3 in which one dimension of the transform data set represents time.

8. A method of data compression as claimed in any one of claims 1 to 3 in which the transform is a frequency-separating transform.

9. A method of data compression as claimed in claim 8 in which the transform decorrelates at least part of the transform data set better in one direction than in a perpendicular direction, and in which a first algorithm is applied by carrying out a one-dimensional scan in the direction of greatest correlation.

10. A method of data compression as claimed in claim 1 in which the transform is two-dimensional.

11. A method of data compression as claimed in claim 1 in which the algorithms are applied by sequential one-dimensional scans through the data.

12. A method of data compression as claimed in claim 11 in which the scans successively switch between directions within the data.

13. A method of data compression as claimed in claim 11 in which successive scans continue in the same direction until an atom is located of lower magnitude than atoms which have previously been located in scans in other directions, and in which the scan direction is then changed.

14. A method of data compression as claimed in claim 13 in which the scan direction is changed to that direction in which an atom of highest current magnitude as previously been located.

15. A method as claimed in claim 2 including applying a function map to the convolved data sets before determining the location of greatest magnitude.

16. A method as claimed in claim 15 in which the function map represents a sensory model.

17. A method as claimed in claim 15 in which the function map represents a psychoacoustic model.

18. A method as claimed in claim 15 in which the function map represents a psychovisual model.

19. A method as claimed in any one of claims 15 to 18 in which the function map is multiplicatively applied.

20. A method as claimed in any one of claims 15 to 19 in which the function map is additively or subtractively applied.

21. A method as claimed in claim 2 in which the second one-dimensional basis functions extend in the spatial domain.

22. A method as claimed in claim 2 in which the second one-dimensional basis functions extend in the time domain.

23. A method as claimed in claim 2, including the additional steps of:

(a) convolving the transform data at the said location with each of a plurality of third one-dimensional basis functions; and

(b) determining a third basis function of a greatest magnitude;

and in which the atom is further derived from the third basis function corresponding to the greatest determined magnitude.

24. A method as claimed in claim 2, in which the second basis function representative of the greatest magnitude is determined without further searching in the region of the said location.

25. A method as claimed in claim 2, in which the second basis function representative of the greatest magnitude is determined at least partly by searching a local area in the region of the said location.

26. A method of data compression comprising:

(a) applying a transform to multi-dimensional data to generate a multi-dimensional transform data set;

(b) convolving the transform data set with each of a plurality of first one-dimensional basis functions to generate a corresponding plurality of convolved data sets;

(c) determining a first location in a first direction across all the convolved data sets, and a first basis function representative of a greatest magnitude; and representing part of the transform data surrounding the first location with a first atom derived from the first function corresponding to the greatest determined magnitude;

(d) subtracting the first atom from the transform data set to create a new data set;

(e) convolving the new data set with each of a plurality of second one-dimensional basis functions;

(f) determining a second location in a second direction across all the convolved data sets, and a second basis function representative of a greatest magnitude; and representing part of the new data set

surrounding a second location with a second atom derived from the second function corresponding to the greatest determined magnitude;

(g) subtracting the second atom from the new data set to create a further new data set;

(h) repeating step (b) with the further new data set, and then re-applying steps (c) to (f); and

(i) outputting as quantized transform data coded versions of the atoms derived at steps (c) and (f).

27. A method of data compression as claimed in claim 26 in which the first location and the second location are coincident.

28. A coder for data compression comprising means for applying a transform to time-varying data to generate a multi-dimensional transform data set, and a coder for coding the transform data set by applying a plurality of one-dimensional matching pursuits algorithms, one for each dimension.

29. A coder for data compression comprising:

(a) means for applying a transform to multi-dimensional data to generate a multi-dimensional transform data set;

(b) means for convolving the transform data set with each of a plurality of first one-dimensional basis functions to generate a corresponding plurality of convolved data sets;

(c) means for determining a location in a first direction across all the convolved data sets, and a first basis function representative of a greatest magnitude;

(d) means for representing part of the transform data surrounding the said location with an atom derived from the first function corresponding to the greatest determined magnitudes;

(g) means for subtracting the atom from the transform data set to create a new data set;

(h) means for repeatedly updating the convolved data sets by convolving any changed part of the transform data set with each of the plurality of first one-dimensional basis functions; and

(i) means for outputting as transform data coded versions of the derived atoms.

30. A coder for data compression as claimed in Claim 29 including:

(c1) means for convolving the transform data at the said location with each of a plurality of second one-dimensional basis functions; and

(c2) means for determining a second basis function representative of a greatest magnitude;

and in which the means for representing part of the transform data further operates upon the second basis functions.

31. A coder for data compression comprising:

(a) means for applying a transform to multi-dimensional data to generate a multi-dimensional transform data set;

(b) means for convolving the transform data set with each of a plurality of first one-dimensional basis functions to generate a corresponding plurality of convolved data sets;

(c) means for determining a first location in a first direction across all the convolved data sets, and a first basis function representative of a greatest magnitude; and representing part of the transform data surrounding the first location with a first atom derived from the first function corresponding to the greatest determined magnitude.

32. A coder for data compression comprising:

(a) means for applying a transform to multi-dimensional data to generate a multi-dimensional transform data set;

(b) means for convolving the transform data set with each of a plurality of first one-dimensional basis functions to generate a corresponding plurality of convolved data sets;

(c) means for determining a first location in a first direction across all the convolved data sets, and a first basis function representative of a greatest magnitude; and representing part of the transform data surrounding the first location with a first atom derived from the first function corresponding to the greatest determined magnitude;

(d) means for subtracting the first atom from the transform data set to create a new data set;

(e) means for convolving the new data set with each of a plurality of second one-dimensional basis functions;

(f) means for determining a second location in a second direction across all the convolved data sets, and a second basis function representative

of a greatest magnitude; and representing part of the new data set surrounding a second location with a second atom derived from the second function corresponding to the greatest determined magnitude;

(g) means for subtracting the second atom from the new data set to create a further new data set;

(h) means for repeating step (b) with the further new data set, and then re-applying steps (c) to (f); and

(i) means for outputting as transform data coded versions of the atoms derived at steps (c) and (f).

33. A codec including a coder as claimed in any one of claims 28 to 32.

34. A computer program for carrying out a method as claimed in any one of claims 1 to 27.

35. A machine-readable data carrier carrying a computer program as claimed in claim 34.

36. A method as claimed in Claim 2 including the further steps of:

(c1) convolving the transform data at the said location with each of a plurality of second one-dimensional basis functions;

(c2) determining a second basis function representative of a greatest magnitude;

and including, at step (d), representing part of the transform data surrounding the said location with an atom derived both from the first and

from the second basis functions corresponding to the greatest determined magnitudes.

37. A method of data compression as claimed in claim 11 in which the sequential one-dimensional scans through the data are orthogonal.

Abstract

A method and apparatus for data compression comprises applying a decorrelating transform to multi-dimensional data to be compressed, then using a sequence of one or more one-dimensional matching pursuits algorithms to code the output of the transform. The invention finds particular application in video and still image coders, particularly real-time coders, those intended for use at very low bit rates, and those for which scalable bit rates are of importance.

(Figure 2)

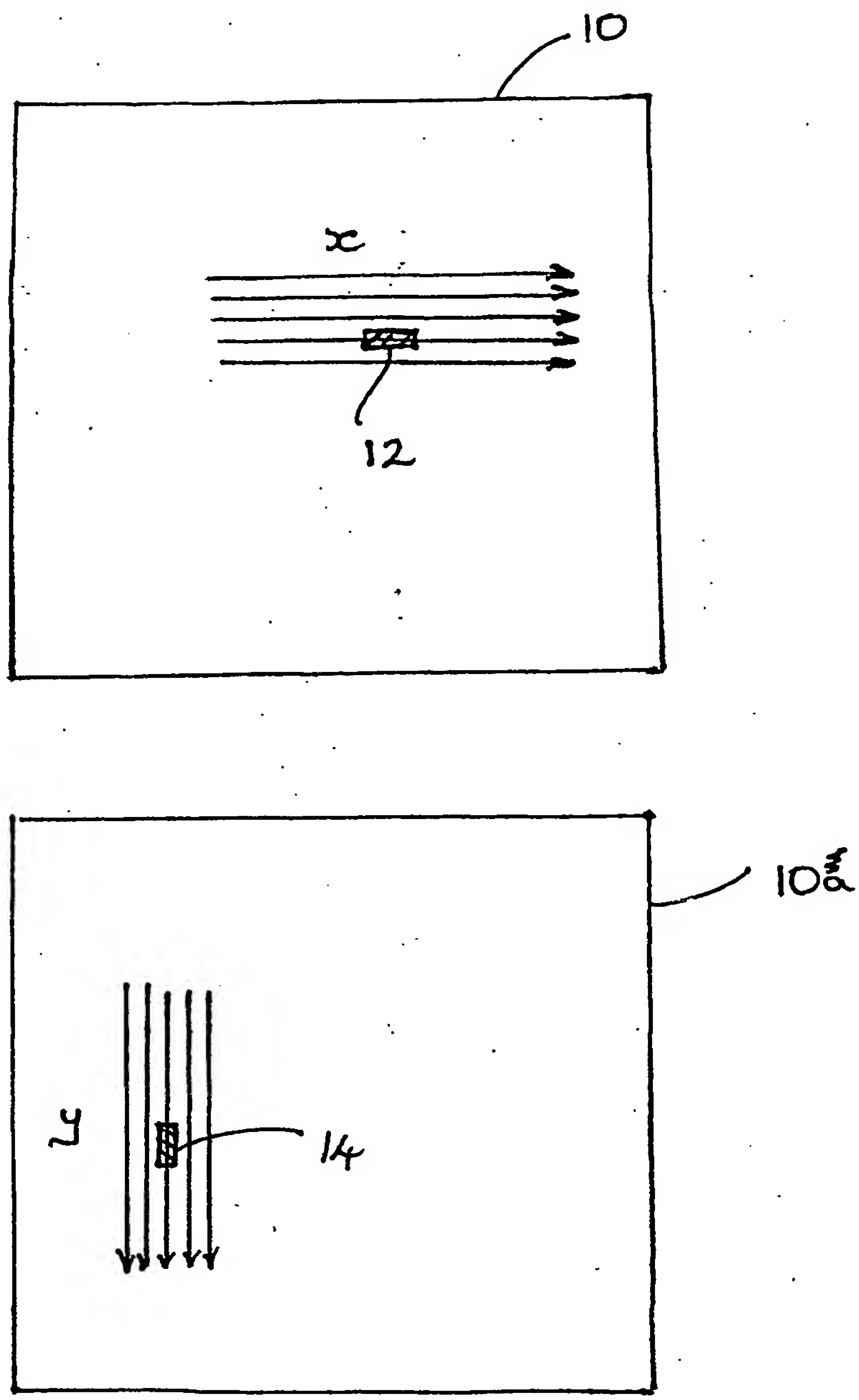


Figure 1

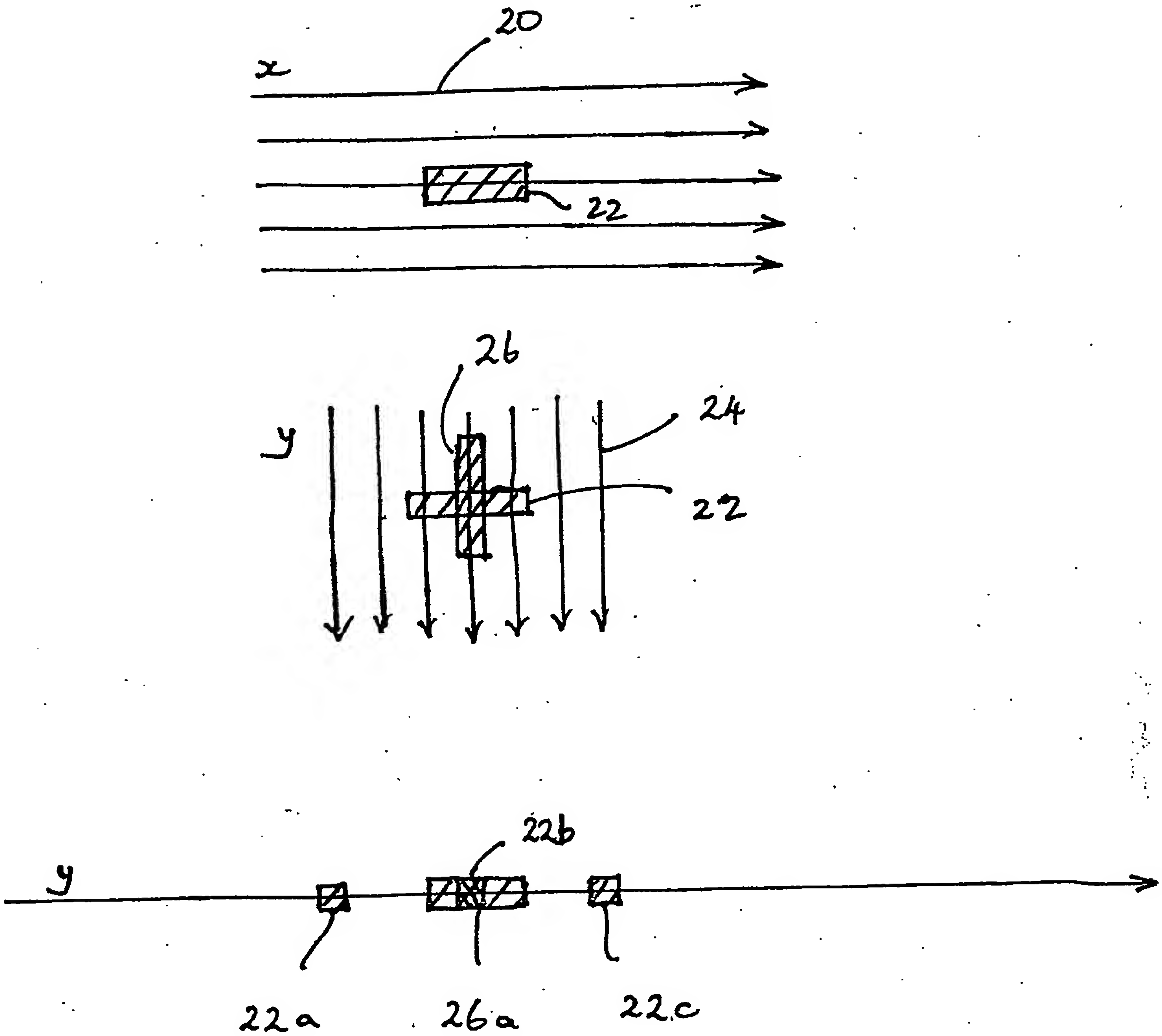


Figure 2

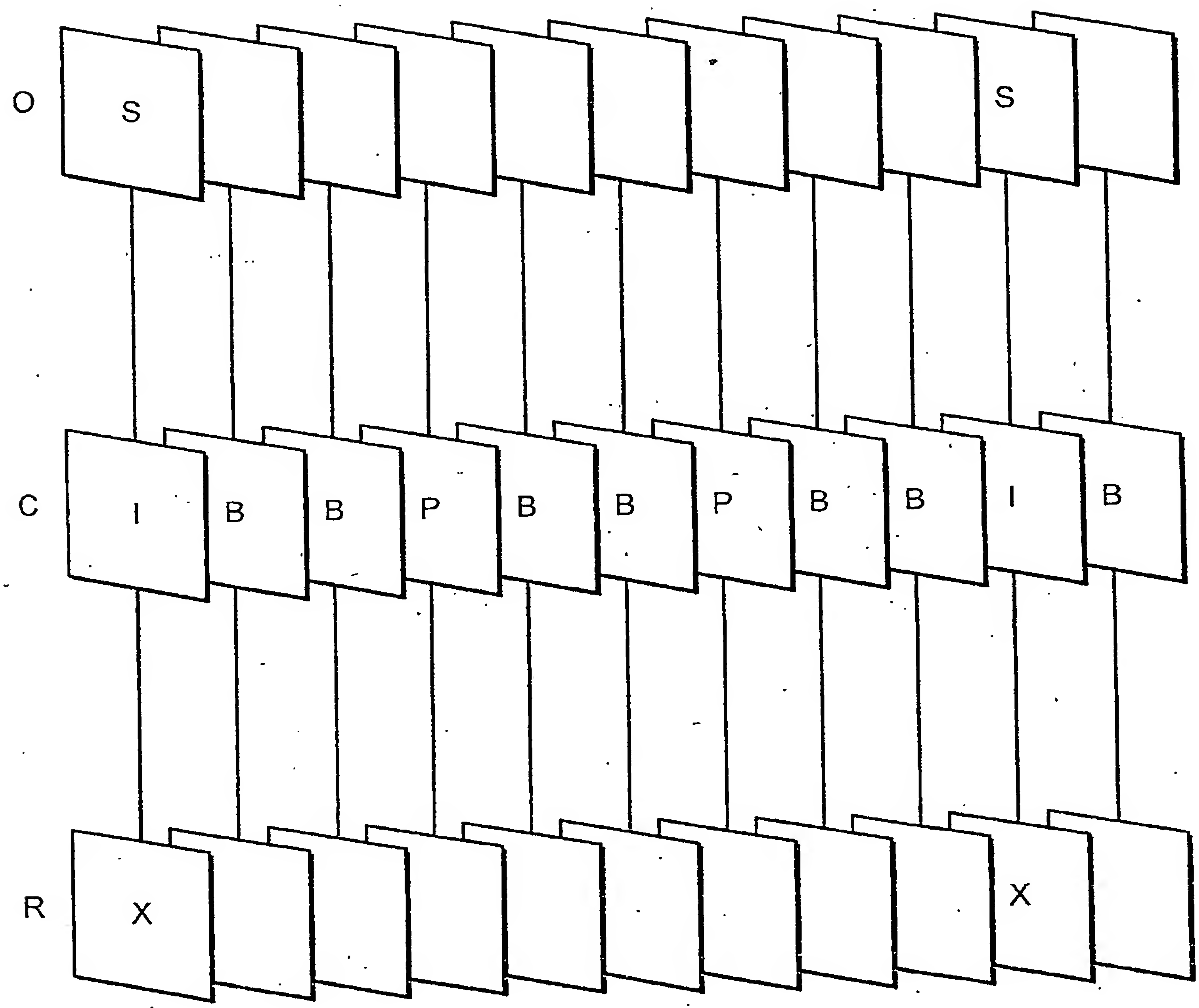


FIG.3

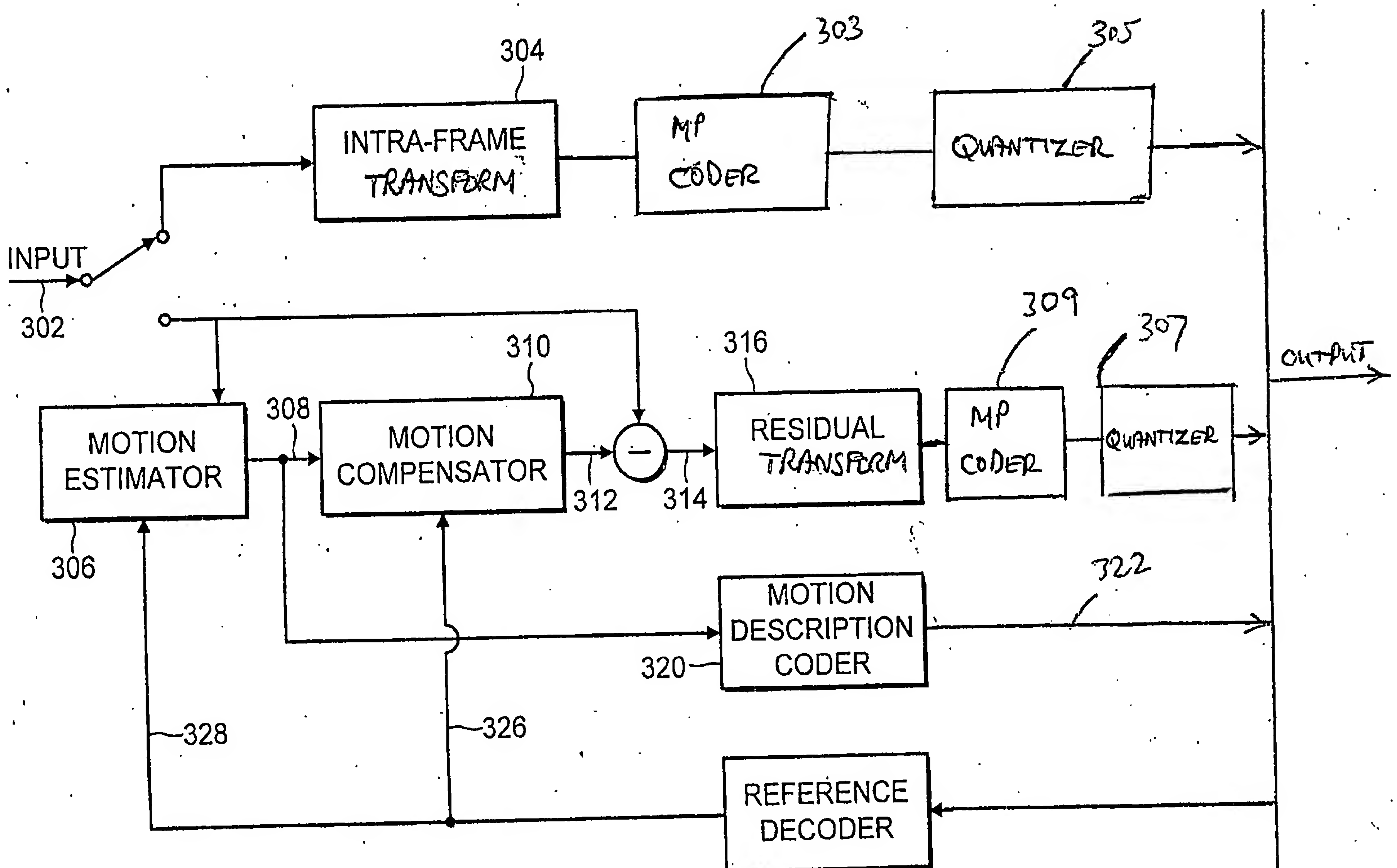

INTRA-FRAME TRANSFORM
304
MP CODER
303
QUANTIZER
305
INPUT
302
MOTION ESTIMATOR
306
308
MOTION COMPENSATOR
310
312
314
RESIDUAL TRANSFORM
316
MP CODER
309
QUANTIZER
307
OUTPUT
MOTION DESCRIPTION CODER
320
322
328
326
REFERENCE DECODER
324

FIG. 4

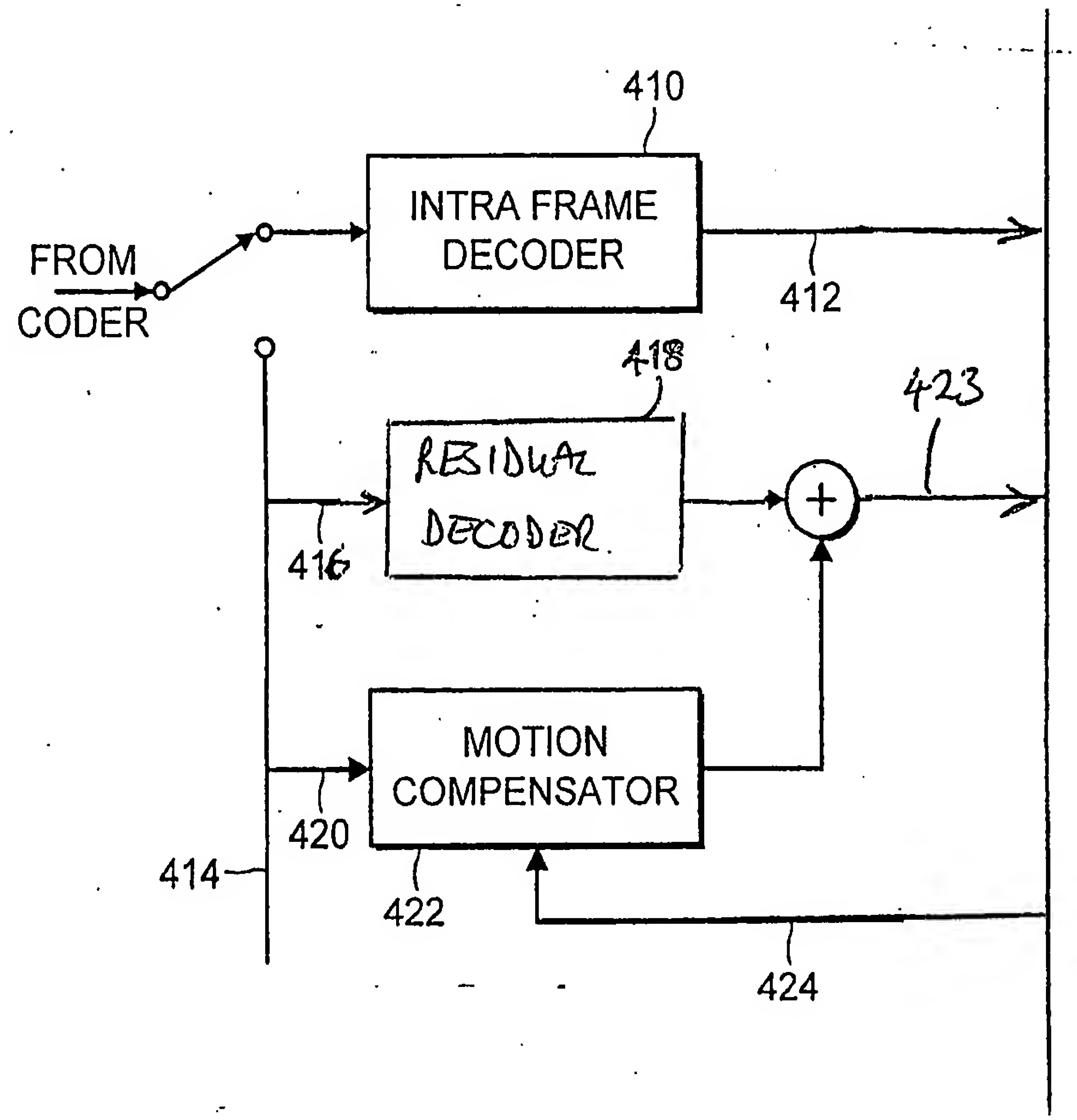
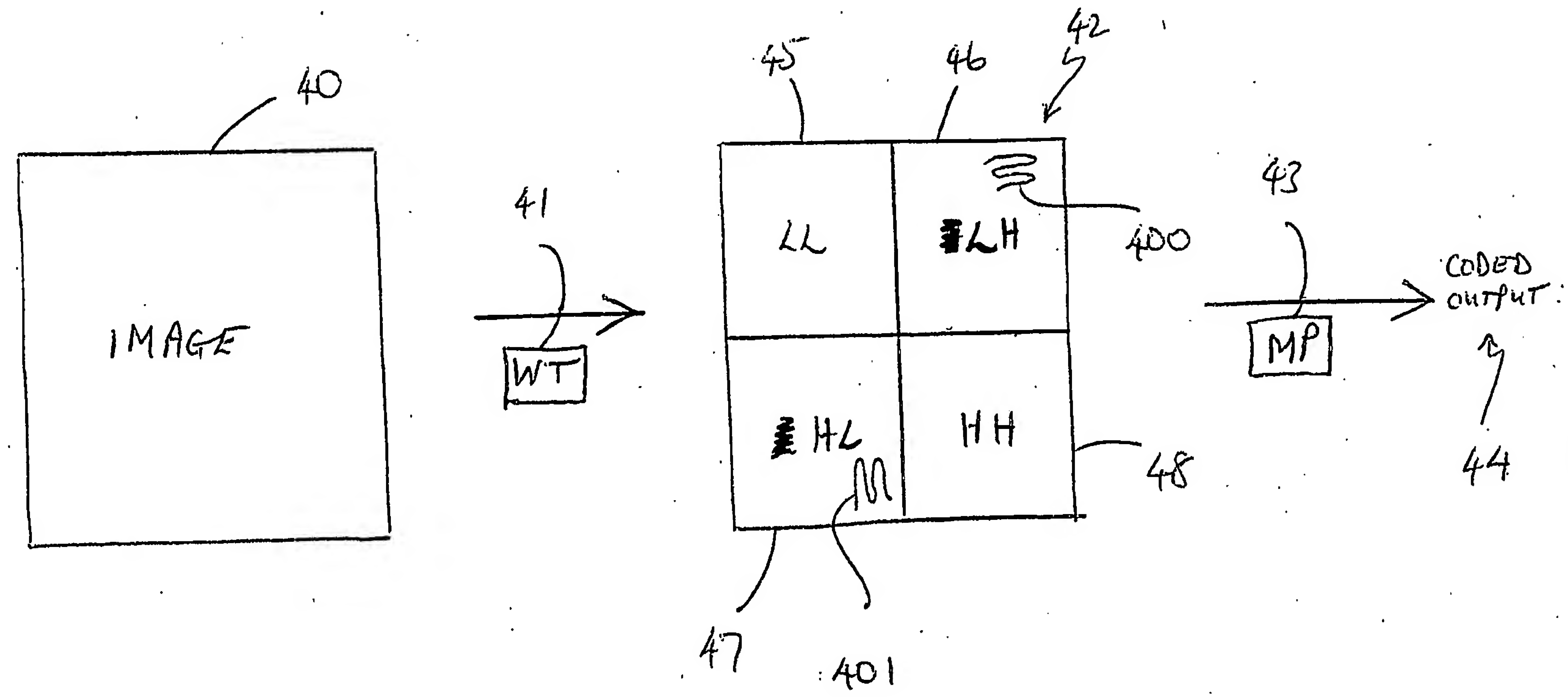

410
INTRA FRAME DECODER
FROM CODER
412
418
RESIDUAL DECODER
423
416
MOTION COMPENSATOR
420
414
422
424

Figure 5



40
IMAGE
41
WT
45
46
42
LL
LH
400
HL
HH
48
47
401
43
MP
CODED OUTPUT
44

Figure 6

PCT/US2005/000612